

The Computer Program DISCCHSE

This program fits the discrete choice model to data using the principle of maximum likelihood. The situation is that one or more animals makes S independent choices of which resource unit to use. At the i th choice there are n_i units that can be selected, in what is called the choice set. Only one of these units is chosen. Each unit is described by its values for p variables X_1, X_2, \dots, X_p , and the probability that the j th unit is selected at the i th choice is

$$p_{ij} = \exp(\beta_1 x_{ij1} + \beta_2 x_{ij2} + \dots + \beta_p x_{ijp}) / \sum_{k=1}^{n_i} \exp(\beta_1 x_{ik1} + \beta_2 x_{ik2} + \dots + \beta_p x_{ikp}),$$

where x_{iku} is the value of X_u for the k th unit that is available at the i th choice.

Data Input

The data takes the following format:

- Line 1: A title for the data set.
Line 2: NUNIT, NX, where NUNIT is the total number of units for which data are available in all the choice sets, and NX is the number of X variables measured on units (called p above)..
Line 3: IND(1), SET(1), X(1,1), X(1,2), ..., X(1,NX), where $X(1) = 1$ if this unit is selected, or otherwise is 0, SET(1) is the number of the choice set that this unit is in, and X(1,1) to X(1,NX) are the values of the X variables (in order) for this unit.
. . .
Line NUNIT+2: IND(NUNIT), SET(NUNIT), X(NUNIT,1), X(NUNIT,2), ..., X(NUNIT,NX), where the terms have the same meaning as on line 3, but for the last resource unit in the data set.

Note that the number of choice sets in total (S above) is SET(NUNIT) because the results must be given for choice set 1, then choice set 2, and so on up to the last choice set. In other words, SET(1) = 1, SET(2) = 2, etc. Also, the number of units in each choice set can vary.

An example set of test data for the program is as follows, where there are 250 resource units altogether in 50 choice sets with five units in each set, and where two X variables are measured on each resource unit:

Test Data for Discrete Choice

250	2		
0	1	4.7	0.7
0	1	0.7	0.2
0	1	3.8	0.1
1	1	7.4	0.1
0	1	2.6	0.8
0	2	0.5	4.8
0	2	4.7	5.4
0	2	2.1	3.4
1	2	8.9	1.1
0	2	9.6	6.2
0	3	3.7	6.2
1	3	7.7	7.0
0	3	4.6	5.3
0	3	2.2	1.3
0	3	1.4	6.5
0	4	1.6	6.8
0	4	2.0	5.8
1	4	8.8	1.4
0	4	5.5	6.9
0	4	1.2	0.0
0	5	5.6	8.8
1	5	5.0	1.9
0	5	7.3	3.6
0	5	6.0	6.4
0	5	2.1	4.0
0	6	1.0	8.6
0	6	3.2	5.6
1	6	8.5	8.8
0	6	6.0	5.5
0	6	0.9	8.6
0	7	8.5	8.5
1	7	5.3	0.7
0	7	2.0	2.7
0	7	4.6	5.3
0	7	3.9	4.6
0	8	4.6	5.6
1	8	4.1	2.5
0	8	1.9	2.9
0	8	3.0	2.7
0	8	2.3	7.3
0	9	3.0	4.6
0	9	1.7	7.7
0	9	6.6	9.4
0	9	2.3	8.7
1	9	5.7	7.4
1	10	7.3	3.4
0	10	0.1	9.9
0	10	3.6	5.5
0	10	5.8	6.1
0	10	7.2	8.1
0	11	5.7	2.4
0	11	3.0	5.0
0	11	5.1	2.8
0	11	8.5	7.9
1	11	9.2	3.7
1	12	7.3	1.3
0	12	2.8	4.6
0	12	1.2	6.3
0	12	0.5	1.1
0	12	6.3	5.1
0	13	0.5	3.3
1	13	3.1	0.9
0	13	1.3	9.1
0	13	0.9	2.9
0	13	3.3	2.5
0	14	1.2	9.3
0	14	7.2	8.1
0	14	1.4	6.3
0	14	1.3	3.1
1	14	6.2	3.9
0	15	0.1	0.3

0	15	4.7	0.3
0	15	2.9	4.8
1	15	8.1	2.5
0	15	1.1	3.6
0	16	0.9	7.3
0	16	6.1	6.1
1	16	3.4	0.5
0	16	2.3	3.3
0	16	1.4	5.4
0	17	2.5	4.1
0	17	0.0	5.2
0	17	5.4	6.5
0	17	7.8	8.5
1	17	1.7	0.1
0	18	7.4	8.2
1	18	9.9	0.4
0	18	3.0	7.8
0	18	4.2	8.3
0	18	0.2	1.6
0	19	8.8	1.6
0	19	7.7	4.4
0	19	0.5	4.9
0	19	3.3	7.7
1	19	6.6	0.9
0	20	7.3	2.8
0	20	0.6	7.4
1	20	8.2	8.9
0	20	5.7	5.8
0	20	5.5	9.0
0	21	8.2	3.8
0	21	1.4	8.3
1	21	7.4	2.5
0	21	0.3	5.2
0	21	2.1	1.6
0	22	0.3	1.2
1	22	8.1	4.3
0	22	7.6	6.3
0	22	0.4	4.4
0	22	8.9	8.8
0	23	1.4	6.5
1	23	9.9	4.5
0	23	3.2	6.1
0	23	9.4	3.9
0	23	2.7	7.8
0	24	1.6	5.5
0	24	3.0	8.8
1	24	1.0	2.0
0	24	9.4	7.6
0	24	5.1	2.2
0	25	0.6	0.3
1	25	5.5	0.5
0	25	2.4	5.1
0	25	4.9	2.0
0	25	8.2	7.2
0	26	2.2	3.8
0	26	5.5	2.6
0	26	6.8	3.8
0	26	0.2	8.6
1	26	9.9	6.4
0	27	5.7	5.8
0	27	5.5	9.3
0	27	2.4	4.5
0	27	0.5	0.7
1	27	9.7	8.1
0	28	3.7	3.1
0	28	9.9	6.9
0	28	2.4	2.7
0	28	7.6	9.9
1	28	8.8	3.1
0	29	4.2	7.6
0	29	1.7	9.1
0	29	4.8	4.6
0	29	8.9	9.3

1	29	4.9	0.5
0	30	5.9	4.1
0	30	9.0	3.6
0	30	1.4	0.2
0	30	6.6	9.7
1	30	9.8	6.3
0	31	7.7	7.4
1	31	6.6	3.0
0	31	0.1	3.5
0	31	5.0	2.2
0	31	9.8	8.8
0	32	0.3	8.2
0	32	1.5	8.4
0	32	7.4	8.0
0	32	2.6	3.3
1	32	7.5	7.7
0	33	5.8	0.6
0	33	9.0	5.6
0	33	1.8	1.3
0	33	0.1	1.0
1	33	6.4	2.0
0	34	2.8	7.2
0	34	6.3	4.4
0	34	2.0	0.1
1	34	7.9	2.9
0	34	2.6	8.8
0	35	8.4	6.5
0	35	2.0	8.8
0	35	8.6	8.5
1	35	9.2	1.0
0	35	7.5	1.5
0	36	7.4	7.4
0	36	0.0	5.7
0	36	3.2	0.4
1	36	9.9	1.5
0	36	9.8	5.9
0	37	8.0	5.7
1	37	7.2	2.5
0	37	3.8	4.3
0	37	8.1	5.8
0	37	1.1	9.6
0	38	5.9	6.4
0	38	7.6	0.3
1	38	9.2	0.3
0	38	9.5	9.6
0	38	3.1	9.8
1	39	7.9	1.3
0	39	8.0	5.1
0	39	4.9	9.8
0	39	0.1	5.2
0	39	1.2	8.5
1	40	9.6	0.8
0	40	7.4	9.8
0	40	1.6	4.4
0	40	4.9	1.1
0	40	5.5	7.6
0	41	8.3	9.8
0	41	5.7	1.9
0	41	0.0	3.9
0	41	9.3	5.1
1	41	7.4	0.8
0	42	0.5	7.2
0	42	2.5	5.6
1	42	2.1	0.2
0	42	7.8	7.5
0	42	3.1	6.2
0	43	3.6	0.9
0	43	1.6	8.0
0	43	7.8	4.0
0	43	5.6	7.6
1	43	6.1	2.2
0	44	4.3	9.9
0	44	6.5	2.6

0	44	6.7	6.1
0	44	5.3	8.0
1	44	9.9	2.3
0	45	5.8	7.6
1	45	9.6	3.2
0	45	1.4	5.1
0	45	9.6	8.7
0	45	7.6	4.8
0	46	6.9	7.7
0	46	3.2	2.5
0	46	4.5	4.8
1	46	7.6	2.9
0	46	5.5	2.4
0	47	1.2	3.7
0	47	1.0	7.7
1	47	1.9	3.1
0	47	1.7	6.8
0	47	6.9	9.4
0	48	5.3	6.8
0	48	0.4	2.9
0	48	0.4	4.0
0	48	8.6	9.2
1	48	9.0	1.7
0	49	5.3	6.6
0	49	7.8	2.7
0	49	4.1	0.8
1	49	5.7	1.0
0	49	7.9	8.6
0	50	2.2	4.0
1	50	8.6	2.0
0	50	4.8	5.5
0	50	0.5	6.7
0	50	2.7	5.8

Computer Output

The output obtained from the program with the above data are shown below. Comments are provided in italics on the right.

```

*****
*                               DISCCHSE                               *
*                               *                                     *
* A program for fitting the discrete choice model to data.          *
* This program is based on the program MAXLIK written by           *
* T.E. Reed for fitting a general multinomial model.               *
*                               *                                     *
* There can be up to 5000 units, with up to 50 estimated          *
* parameters.                                                       *
*                               *                                     *
*                               Written by                            *
*                               Bryan F.J. Manly                     *
*                               Western EcoSystems Technology Inc.    *
*                               Cheyenne, Wyoming                    *
*                               bmanly@west-inc.com                  *
*                               *                                     *
*                               Version 1.0, Dated July 2001        *
*****

```

Date: 28/ 5/2002 Time: 13:41

```

Data Title: Test Data for Discrete Choice
Number of Units = 250
Unit Use Set X-Values
1 0 1 4.70 0.70
2 0 1 0.70 0.20
3 0 1 3.80 0.10
4 1 1 7.40 0.10

```

5	0	1	2.60	0.80
6	0	2	0.50	4.80
7	0	2	4.70	5.40
8	0	2	2.10	3.40
9	1	2	8.90	1.10
10	0	2	9.60	6.20
11	0	3	3.70	6.20
12	1	3	7.70	7.00
13	0	3	4.60	5.30
14	0	3	2.20	1.30
15	0	3	1.40	6.50
16	0	4	1.60	6.80
17	0	4	2.00	5.80
18	1	4	8.80	1.40
19	0	4	5.50	6.90
20	0	4	1.20	0.00
21	0	5	5.60	8.80
22	1	5	5.00	1.90
23	0	5	7.30	3.60
24	0	5	6.00	6.40
25	0	5	2.10	4.00
26	0	6	1.00	8.60
27	0	6	3.20	5.60
28	1	6	8.50	8.80
29	0	6	6.00	5.50
30	0	6	0.90	8.60
31	0	7	8.50	8.50
32	1	7	5.30	0.70
33	0	7	2.00	2.70
34	0	7	4.60	5.30
35	0	7	3.90	4.60
36	0	8	4.60	5.60
37	1	8	4.10	2.50
38	0	8	1.90	2.90
39	0	8	3.00	2.70
40	0	8	2.30	7.30
41	0	9	3.00	4.60
42	0	9	1.70	7.70
43	0	9	6.60	9.40
44	0	9	2.30	8.70
45	1	9	5.70	7.40
46	1	10	7.30	3.40
47	0	10	0.10	9.90
48	0	10	3.60	5.50
49	0	10	5.80	6.10
50	0	10	7.20	8.10
.
247	1	50	8.60	2.00
248	0	50	4.80	5.50
249	0	50	0.50	6.70
250	0	50	2.70	5.80

*** NULL MODEL FIT ***

The null model has no parameters and gives all choices the same probability

Log-likelihood = -80.472

Deviance = AIC = 160.944 with 250 df

Models allowing for selection related to the X-variables are now fitted. This process can be repeated as much as necessary. First, variables to be used are chosen.

Variables chosen for use are:

1

Just fitting variable 1 first.

*** SELECTION MODEL FIT ***

Parameters of the fitted model are the coefficients of the 1 X-variables.

Summary of Iterations

Iter	Initial log-lik	Final log-lik	Halvings
1	-80.472	-60.504	1

2	-60.504	-59.680	1
3	-59.680	-59.669	1
4	-59.669	-59.669	2

Final Deviance = 119.34 with 249 df
AIC = 121.34

Param	Final estimate	Standard error	Last change
1	0.4030	0.0753	0.0002

Correlation matrix for parameters
1 1.000

Now asked for all (two) variables to be included in the model.

*** SELECTION MODEL FIT ***

Parameters of the fitted model are the coefficients of the 2 X-variables.

Summary of Iterations

Iter	Initial log-lik	Final log-lik	Halvings
1	-80.472	-37.943	1
2	-37.943	-29.951	1
3	-29.951	-27.445	1
4	-27.445	-27.028	1
5	-27.028	-27.010	1
6	-27.010	-27.010	1
7	-27.010	-27.010	2

Final Deviance = 54.02 with 248 df
AIC = 58.02

Param	Final estimate	Standard error	Last change
1	0.9377	0.1992	-0.0006
2	-0.8611	0.1756	0.0004

Correlation matrix for parameters
1 1.000
2 -0.806 1.000

The true parameter values for these artificial data are +1 and -1, so the estimates are quite reasonable at 0.9377 and -0.8611.

Bryan F.J. Manly
Western EcoSystems Technology Inc.
bmanly@west-inc.com